

CLAIMS

1. A method of training a computer classification system which can be defined by a network comprising a number of n-tuples or Look Up Tables (LUTs),
5 with each n-tuple or LUT comprising a number of rows corresponding to at least a subset of possible classes and further comprising a number of columns being addressed by signals or elements of sampled training input data examples, each column being defined by a vector having cells with values, wherein
the column vector cell values are determined based on one or more
10 training sets of input data examples for different classes so that at least part of the cells comprise or point to information based on the number of times the corresponding cell address is sampled from one or more sets of training input examples, said method being characterised in that
one or more output score functions are determined for evaluation of at least
15 one output score value per class, and
one or more decision rules are determined to be used in combination with at least part of the obtained output scores to determine a winning class, wherein said determination of the output score functions and decision rules comprises
determining output score functions based on the information of at least part
20 of the determined column vector cell values, and adjusting at least part of the output score functions based on an information measure evaluation, and/or
determining decision rules based on the information of at least part of the determined column vector cell values, and adjusting at least part of the decision rules based on an information measure evaluation.
2. A method according to claim 1, wherein the output score functions are determined based on a validation set of input data examples.
3. A method according to claim 1 or 2, wherein the decision rules are
30 determined based on a validation set of input data examples.
4. A method according to any of the claims 1-3, wherein determination of the output score functions is based on an information measure evaluating the performance on the validation example set.

35

5. A method according to any of the claims 1-4, wherein determination of the decision rules is based on an information measure evaluating the performance on the validation example set.
- 5 6. A method according to any of the claims 3-5, wherein the validation example set equals at least part of the training set and the information measure is based on a leave-one-out cross validation evaluation.
- 10 7. A method according to any of the claims 3-6, wherein the validation set comprises at least part of the training set(s) of input data examples.
8. A method according to any of the claims 1-7, wherein the output score functions are determined by a set of parameter values.
- 15 9. A method according to any of the claims 1-8, wherein determination of the output score functions comprises initialising the output score functions.
10. A method according to claim 9, wherein the initialisation of the output score functions comprises determining a number of set-up parameters.
- 20 11. A method according to claims 9 or 10, wherein the initialisation of the output score functions comprises setting all output score functions to a pre-determined mapping function.
- 25 12. A method according to any of the claims 1-11, wherein determination of the decision rules comprises initialising the decision rules.
13. A method according to claim 12, wherein the initialisation of the decision rules comprises setting the rules to a pre-determined decision scheme.
- 30 14. A method according to any of the claims 10-13, wherein the adjustment comprises changing the values of the set-up parameters.
- 35 15. A method according to any of the claims 1-14, wherein the determination of the column vector cell values comprises the training steps of

- 5
- a) applying a training input data example of a known class to the classification network, thereby addressing one or more column vectors,
 - b) incrementing, preferably by one, the value or vote of the cells of the addressed column vector(s) corresponding to the row(s) of the known class, and
 - c) repeating steps (a)-(b) until all training examples have been applied to the network.
- 10
16. A method according to any of the claims 1-15, wherein the adjustment process comprises the steps of
- determining a global quality value based on at least part of the column vector cell values,
 - determining if the global quality value fulfils a required quality criterion,
- 15
- and
- adjusting at least part of the output score functions until the global quality criterion is fulfilled.
17. A method according to any of the claims 1-16, wherein the adjustment process comprises the steps of
- 20
- a) selecting an input example from the validation set(s),
 - b) determining a local quality value corresponding to the sampled validation input example, the local quality value being a function of at least part of the addressed column cell values,
 - c) determining if the local quality value fulfils a required local quality criterion, if not,
- 25
- adjusting one or more of the output score functions if the local quality criterion is not fulfilled,
 - d) selecting a new input example from a predetermined number of examples of the validation set(s),
- 30
- e) repeating the local quality test steps (b)-(d) for all the predetermined validation input examples,
 - f) determining a global quality value based on at least part of the column vectors being addressed during the local quality test,
 - g) determining if the global quality value fulfils a required global quality criterion,
- 35
- and,

3/
4

h) repeating steps (a)-(g) until the global quality criterion is fulfilled.

18. A method according to any of the claims 1-17, wherein the adjustment process comprises the steps of

- 5 determining a global quality value based on at least part of the column vector cell values,
determining if the global quality value fulfils a required quality criterion,
and
adjusting at least part of the decision rules until the global quality criterion is
10 fulfilled.

19. A method according to any of the claims 1-18, wherein the adjustment process comprises the steps of

- a) selecting an input example from the validation set(s),
15 b) determining a local quality value corresponding to the sampled validation input example, the local quality value being a function of at least part of the addressed column cell values,
c) determining if the local quality value fulfils a required local quality criterion, if not,
20 adjusting one or more of the decision rules if the local quality criterion is not fulfilled,
d) selecting a new input example from a predetermined number of examples of the validation set(s),
e) repeating the local quality test steps (b)-(d) for all the predetermined valida-
25 tion input examples,
f) determining a global quality value based on at least part of the column vectors being addressed during the local quality test,
g) determining if the global quality value fulfils a required global quality criterion, and,
30 h) repeating steps (a)-(g) until the global quality criterion is fulfilled.

20. A method according to any of the claims 1-15, wherein the adjustment process comprises the steps of

- 35 determining a global quality value based on at least part of the column vector cell values,

determining if the global quality value fulfils a required quality criterion,
and
adjusting at least part of the output score functions and part of the decision
rules until the global quality criterion is fulfilled.

5

21. A method according to any of the claims 1-15 or 20, wherein the adjustment process comprises the steps of

- a) selecting an input example from the validation set(s),
- b) determining a local quality value corresponding to the sampled validation
input example, the local quality value being a function of at least part of the
addressed column cell values,
- c) determining if the local quality value fulfils a required local quality criterion, if
not,
adjusting one or more of the output score functions and the decision rules if
the local quality criterion is not fulfilled,
- d) selecting a new input example from a predetermined number of examples of
the validation set(s),
- e) repeating the local quality test steps (b)-(d) for all the predetermined validation
input examples,
- f) determining a global quality value based on at least part of the column vectors
being addressed during the local quality test,
- g) determining if the global quality value fulfils a required global quality criterion,
and,
- h) repeating steps (a)-(g) until the global quality criterion is fulfilled.

25

22. A method according to claim 17, 19 or 21, wherein steps (b)-(d) are
carried out for all examples of the validation set(s).

30 23. A method according to any of the claims 16-22, wherein the local
and/or global quality value is defined as functions of at least part of the column cells.

24. A method according to any of the claims 16-23, wherein the adjustment
iteration process is stopped if the quality criterion is not fulfilled after a given
number of iterations.

35

25. A method of classifying input data examples into at least one of a plurality of classes using a computer classification system configured according to any of the claims 1-24, whereby column cell values for each n-tuple or LUT and output score functions and/or decision rules are determined using on one or more training or validation sets of input data examples, said method comprising

- a) applying an input data example to be classified to the configured classification network thereby addressing column vectors in the set of n-tuples or LUTs,
- b) selecting a set of classes which are to be compared using a given set of output score functions and decision rules thereby addressing specific rows in the set of n-tuples or LUTs,
- c) determining output score values as a function of the column vector cells and using the determined output score functions,
- d) comparing the calculated output values using the determined decision rules, and
- e) selecting the class or classes that win(s) according to the decision rules.

26. A system for training a computer classification system which can be defined by a network comprising a stored number of n-tuples or Look Up Tables (LUTs), with each n-tuple or LUT comprising a number of rows corresponding to at least a subset of possible classes and further comprising a number of columns being addressed by signals or elements of sampled training input data examples, each column being defined by a vector having cells with values, said system comprising

- a) input means for receiving training input data examples of known classes,
- b) means for sampling the received input data examples and addressing column vectors in the stored set of n-tuples or LUTs,
- c) means for addressing specific rows in the set of n-tuples or LUTs, said rows corresponding to a known class,
- d) storage means for storing determined n-tuples or LUTs,
- e) means for determining column vector cell values so as to comprise or point to information based on the number of times the corresponding cell address is sampled from the training set(s) of input examples, characterised in that said system further comprises

34
7

f) means for determining one or more output score functions and one or more decision rules, wherein said output score functions and decision rules determining means is adapted for
determining said output score functions based on the information of at least
5 part of the determined column vector cell values and a validation set of input data examples of known classes, and
determining said decision rules based on the information of at least part of the determined column vector cell values and a validation set of input data examples of known classes, and wherein the means for determining the out-
10 put score functions and decision rules comprises
means for initialising one or more sets of output score functions and/or decision rules, and
means for adjusting output score functions and decision rules by use of at least part of the validation set of input examples.

15 27. A system according to claim 26, wherein the means for determining the output score functions is adapted to determine such functions from a family of output score functions determined by a set of parameter values.

20 28. A system according to claim 26 or 27, wherein said validation set comprises at least part of the training set(s) used for determining the column cell values.

25 29. A system according to any of the claims 26-28, wherein the means for determining the column vector cell values is adapted to determine these values as a function of the number of times the corresponding cell address is sampled from the set(s) of training input examples.

30 30. A system according to any of the claims 26-29, wherein, when a training input data example belonging to a known class is applied to the classification network thereby addressing one or more column vectors, the means for determining the column vector cell values is adapted to increment the value or vote of the cells of the addressed column vector(s) corresponding to the row(s) of the known class, said value preferably being incremented by one.

31. A system according to any of the claims 26-30, wherein the means for adjusting output score functions is adapted to

determine a global quality value based on at least part of column vector cell values,

5 determine if the global quality value fulfils a required global quality criterion, and

adjust at least part of the output score functions until the global quality criterion is fulfilled.

10 32. A system according to any of the claims 26-31, wherein the means for adjusting output score functions and decision rules is adapted to

a) determine a local quality value corresponding to a sampled validation input example, the local quality value being a function of at least part of the addressed vector cell values,

15 b) determine if the local quality value fulfils a required local quality criterion,

c) adjust one or more of the output score functions if the local quality criterion is not fulfilled,

d) repeat the local quality test for a predetermined number of training input examples,

20 e) determine a global quality value based on at least part of the column vectors being addressed during the local quality test,

f) determine if the global quality value fulfils a required global quality criterion, and,

25 g) repeat the local and the global quality test until the global quality criterion is fulfilled.

33. A system according to any of the claims 26-32, wherein the means for adjusting decision rules is adapted to

30 determine a global quality value based on at least part of column vector cell values,

determine if the global quality value fulfils a required global quality criterion, and

adjust at least part of the decision rules until the global quality criterion is fulfilled.

35

36
8

34. A system according to any of the claims 26-33, wherein the means for adjusting output score functions and decision rules is adapted to

- 5 a) determine a local quality value corresponding to a sampled validation input example, the local quality value being a function of at least part of the addressed vector cell values,
- b) determine if the local quality value fulfils a required local quality criterion,
- c) adjust one or more of the decision rules if the local quality criterion is not fulfilled,
- 10 d) repeat the local quality test for a predetermined number of training input examples,
- e) determine a global quality value based on at least part of the column vectors being addressed during the local quality test,
- f) determine if the global quality value fulfils a required global quality criterion, and,
- 15 g) repeat the local and the global quality test until the global quality criterion is fulfilled.

35. A system according to any of the claims 26-30, wherein the means for adjusting decision rules is adapted to

- 20 determine a global quality value based on at least part of column vector cell values,
- determine if the global quality value fulfils a required global quality criterion, and
- adjust least part of the output score functions and decision rules until the
- 25 global quality criterion is fulfilled.

36. A system according to any of the claims 26-30 or 35, wherein the means for adjusting output score functions and decision rules is adapted to

- 30 a) determine a local quality value corresponding to a sampled validation input example, the local quality value being a function of at least part of the addressed vector cell values,
- b) determine if the local quality value fulfils a required local quality criterion,
- c) adjust one or more of the output score functions and decision rules if the local quality criterion is not fulfilled,

37
10

- d) repeat the local quality test for a predetermined number of training input examples,
- e) determine a global quality value based on at least part of the column vectors being addressed during the local quality test,
- 5 f) determine if the global quality value fulfils a required global quality criterion, and,
- g) repeat the local and the global quality test until the global quality criterion is fulfilled.
- 10 37. A system according to any of the claims 31-36, wherein the means for adjusting the output score functions and decision rules is further adapted to stop the iteration process if the global quality criterion is not fulfilled after a given number of iterations.
- 15 38. A system according to any of the claims 26-37, wherein the means for storing n-tuples or LUTs comprises means for storing adjusted output score functions and decision rules and separate means for storing best so far output score functions and decision rules or best so far classification system configuration values.
- 20 39. A system according to claim 38, wherein the means for adjusting the output score functions and decision rules is further adapted to replace previously separately stored best so far output score functions and decision rules with obtained adjusted output score functions and decision rules if the determined global quality value is closer to fulfil the global quality criterion than the global quality value corresponding to previously separately stored best so far output score functions and decision rules.
- 25 40. A system for classifying input data examples of unknown classes into at least one of a plurality of classes, said system comprising:
- 30 storage means for storing a number or set of n-tuples or Look Up Tables (LUTs) with each n-tuple or LUT comprising a number of rows corresponding to at least a subset of the number of possible classes and further comprising a number of column vectors, each column vector being addressed by signals or elements of a sampled input data example, and each column vector hav-

ing cell values being determined during a training process based on one or more sets of training input data examples,
storage means for storing one or more output score functions and/or one or more decision rules, each output score function and/or decision rule being
5 determined during a training or validation process based on one or more sets of validation input data examples, said system further comprising:

input means for receiving an input data example to be classified,
means for sampling the received input data example and addressing column
vectors in the stored set of n-tuples or LUTs,
10 means for addressing specific rows in the set of n-tuples or LUTs, said rows corresponding to a specific class,
means for determining output score values using the stored output score functions and at least part of the stored column vector values, and
means for determining a winning class or classes based on the output score
15 values and stored decision rules.

41. A system according to claim 40, wherein the cell values of the column vectors and the output score functions and/or decision rules of the classification system are determined by use of a training system according to any of the claims
20 26-39.

42. A system according to claim 40, wherein the column vector cell values and the output score functions and/or decision rules are determined during a training process according to any of the claims 1-24.
25